Scientific and Technical Information Center

SEARCH REQUEST FORM

Date: 12/23/99   Requester's Full Name: Amy Nelson   Examiner # : 73840

Art Unit: 1649   Phone (306) 3218   Serial Number: 09/327,230

Results Format Preferred (circle): PAPER   DISK   E-MAIL

****

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: _____

Inventors (please provide full names): Gray et al.

_____

Earliest Priority Date: _____

Search Topic:
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known.

For Sequence Searches Only *Please include all pertinent information (parent, grandchild, divisional, or issued patent numbers) along with the appropriate serial number.*

Please search SEQ ID No. 1.

****

THIS PAGE BLANK (USPTO)

```
******************************************************************

            M P S R C H  (TM)

******************************************************************
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 28 14:30:35 1999; MasPar time 5058.20 Seconds
                1546.464 Million cell updates/sec

Title:
Description: >US-09-327-230-1
                (1-2822) from US09327230.seq
Perfect Score: 2822
N.A. Sequence: 1 gcaaccacacagacaggca.............ttccaggtttgggtttcggg 2822
Comp: cgttgctgtgtgtctgtccgt...........aaggtcaaacccaaagccc

Scoring table: TABLE default
                Gap 6

Nmatch STD :   Dbase 0: Query 0

Searched: 646147 seqs, 138955633 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database: embl58
                1:em_ba1 2:em_ba2 3:em_fun 4:em_htg 5:em_hum1 6:em_hum2
                7:em_in 8:em_om 9:em_or 10:em_ov 11:em_pat 12:em_ph
                13:em_pl 14:em_ro 15:em_sts 16:em_vi

Database: genbank111
                17:gb_ba1 18:gb_ba2 19:gb_htg1 20:gb_htg2 21:gb_in1
                22:gb_in2 23:gb_om 24:gb_ov 25:gb_pat 26:gb_ph 27:gb_pl1
                28:gb_pl2 29:gb_pr1 30:gb_pr2 31:gb_pr3 32:gb_ro
                33:gb_st 34:gb_sts 35:gb_sy 36:gb_un 37:gb_vi

Statistics: Mean 12.024; Variance 6.432; scale 1.869

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

### SUMMARIES

|   | Result No. | Score | % Query Match | Length DB | ID | Description | Pred. No. |
|---|---|---|---|---|---|---|---|
|   | 1 | 2822 | 100.0 | 7129 | 28 | ZMU77346 | Zea mays lethal leaf-s | 0.00e+00 |
|   | 2 | 275 | 9.7 | 5364 | 28 | AF050129 | Zea mays cell wall inv | 5.52e-180 |
| c | 3 | 83 | 2.9 | 7218 | 25 | I66494 | Sequence 14 from paten | 3.13e-35 |
| c | 4 | 75 | 2.7 | 250 | 27 | MZESUT282 | Maize sucrose synthetase | 1.13e-29 |
| c | 5 | 59 | 2.1 | 7218 | 25 | I66494 | Sequence 14 from paten | 5.69e-19 |
| c | 6 | 53 | 1.9 | 10772 | 21 | AF012089 | Drosophila melanogaste | 5.10e-15 |
| c | 7 | 49 | 1.7 | 10772 | 21 | AF012089 | Drosophila melanogaste | 1.71e-12 |
| c | 8 | 45 | 1.6 | 250 | 27 | MZESUT281 | Maize sucrose synthetase | 4.99e-10 |
| c | 9 | 43 | 1.5 | 5057 | 27 | ZMCI31AC3 | z.mays cyp71c3 gene. | 8.02e-09 |
| c | 10 | 43 | 1.5 | 6831 | 27 | ZMA132240 | Zea mays eIF-5 gene, e | 8.02e-09 |
| c | 11 | 39 | 1.4 | 965 | 25 | AR024229 | Sequence 22 from paten | 1.81e-06 |
| c | 12 | 38 | 1.3 | 1056 | 23 | MVU87256 | Mustela vison GT dinuc | 6.78e-06 |
| c | 13 | 37 | 1.3 | 7905 | 27 | MZECPN60B | Corn nuclear-encoded m | 2.51e-05 |
|   | 14 | 35 | 1.2 | 74371 | 31 | AC005369 | Homo sapiens chromosom | 3.27e-04 |
|   | 15 | 30 | 1.1 | 215 | 25 | I28278 | Sequence 5 from patent | 1.48e-01 |
|   | 16 | 31 | 1.1 | 215 | 25 | I28278 | Sequence 5 from patent | 1.48e-01 |
| c | 17 | 32 | 1.1 | 965 | 25 | AR024229 | Sequence 22 from paten | 1.48e-01 |
|   | 18 | 32 | 1.1 | 1056 | 23 | MVU87256 | Mustela vison GT dinuc | 1.36e-02 |
|   | 19 | 31 | 1.1 | 216021 | 31 | HUAC004787 | Homo sapiens Chromosom | 1.36e-02 |
|   | 20 | 32 | 1.1 | 216021 | 31 | HUAC004787 | Homo sapiens Chromosom | 4.54e-02 |
|   | 21 | 28 | 1.1 | 60 | 25 | A62989 | Sequence 1 from Patent | 1.48e-02 |
|   | 22 | 27 | 1.0 | 354 | 21 | OFU09259 | Oxytricha fallax 57kD | 1.48e+00 |
| c | 23 | 28 | 1.0 | 1663 | 23 | MIECORI4 | Yeast mitochondrial or | 1.48e+00 |
|   | 24 | 29 | 1.0 | 7129 | 28 | MVU92534 | Mustela vison microsat | 1.48e-01 |
| c | 25 | 28 | 1.0 | 29482 | 21 | AE001367 | Plasmodium falciparum | 1.48e+00 |
| c | 26 | 29 | 1.0 | 13594 | 22 | CEW0OB5 | Caenorhabditis elegans | 1.48e+00 |
| c | 27 | 28 | 1.0 | 7129 | 28 | ZMU77346 | Zea mays lethal leaf-s | 1.48e+00 |
| c | 28 | 29 | 1.0 | 93861 | 28 | ATAC005167 | Arabidopsis thaliana c | 4.73e-01 |
| c | 29 | 29 | 1.0 | 93578 | 30 | CEY41C4A | Caenorhabditis elegans | 4.73e-01 |
| c | 30 | 29 | 1.0 | 36524 | 21 | AC005731 | Arabidopsis thaliana c | 4.73e-01 |
|   | 31 | 29 | 1.0 | 42366 | 21 | CEF25H2 | Caenorhabditis elegans | 4.73e-01 |
|   | 32 | 28 | 1.0 | 55328 | 28 | PWU02970 | Prototheca wickerhamii | 4.73e-01 |
|   | 33 | 27 | 1.0 | 74371 | 31 | AC005369 | Homo sapiens chromosom | 4.73e-01 |
|   | 34 | 28 | 1.0 | 89138 | 28 | ATAC006069 | Arabidopsis thaliana c | 4.73e-01 |
|   | 35 | 28 | 1.0 | 1466560 | 19 | HSAC006069 | Human PAC clone DJ073F | 4.73e-01 |
|   | 36 | 29 | 1.0 | 143944 | 31 | ATAC005167 | Arabidopsis thaliana c | 4.73e-01 |
|   | 37 | 28 | 1.0 | 1465642 | 19 | CEY41C4A | Caenorhabditis elegans | 4.73e-01 |
|   | 38 | 27 | 1.0 | 152593 | 21 | AC004998 | Human Xq28 BAC RPCI11- | 4.73e-01 |
|   | 39 | 28 | 1.0 | 153098 | 21 | HS580G13 | Human DNA sequence *** | 4.73e-01 |
|   | 40 | 29 | 1.0 | 152593 | 28 | ATAC002332 | Arabidopsis thaliana c | 4.73e-01 |
|   | 41 | 28 | 1.0 | 156152 | 19 | PFMAL3P2 | Plasmodium falciparum | 4.73e-01 |
|   | 42 | 28 | 1.0 | 156151 | 19 | CEY26D4 | Caenorhabditis elegans | 4.73e-01 |
|   | 43 | 29 | 1.0 | 257817 | 19 | AC006909 | Caenorhabditis elegans | 4.73e-01 |
| c | 44 | 29 | 1.0 | 257817 | 19 | AC006909 | Caenorhabditis elegans | 4.73e-01 |
|   | 45 | 28 | 1.0 | 296631 | 20 | AC005308 | Plasmodium falciparum | 1.48e+00 |

*QH573.C38*

### ALIGNMENTS

RESULT 1

```
LOCUS       ZMU77346     7129 bp    DNA           PLN       15-APR-1997
DEFINITION  Zea mays lethal leaf-spot 1 (lls1) gene, partial cds.
ACCESSION   U77346
NID         g1935910
VERSION     U77346.1  GI:1935910
KEYWORDS    .
SOURCE      maize.
  ORGANISM  Zea mays
            Eukaryota; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; vascular plants; seed
            plants; Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
REFERENCE   1  (bases 1 to 7129)
  AUTHORS   Gray,J., Close,P.S., Briggs,S.P. and Johal,G.S.
  TITLE     A novel suppressor of cell death in plants encoded by the Lls1 gene
            of maize
  JOURNAL   Cell 89 (1), 25-31 (1997)
  MEDLINE   97248483
REFERENCE   2  (bases 1 to 7129)
  AUTHORS   Gray,J. and Johal,G.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-1996) Agronomy, University of Missouri, 205
            Curtis Hall, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1..7129
                     /organism="Zea mays"
                     /strain="B73 inbred line"
                     /db_xref="taxon:4577"
                     /chromosome="1S"
                     /map="Bin 1.01 - 1.03"
     CDS             complement(<85..>290)
                     /note="similar to Nicotiana tabacum auxin inducible
                     protein"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAC49677.1"
                     /db_xref="PID:g1935911"
```

```
exon                    /db_xref="GI:1935911"
                        /translation="AIKQLSREQVQVATKFEIQRDESGTRIVCGRPEYVRACCEASLR
                        RLGIDCIDLIYQHHIDTTIPIEETV"
                        3115..3764
                        /note="putative"

mRNA                    /number=1
                        join(3115..3764,3854..4089,4178..4304,5480..5616,
                        5729..5588,6119..6397,6923..7129)
                        /gene="lls-1"

gene                    3396..7129
                        /gene="lls-1"

CDS                     join(3396..3764,3854..4089,4178..4304,5480..5616,
                        5729..5588,6119..6397,6923..>7129)
                        /gene="lls-1"
                        /note="similar to bacterial ring-hydroxylating
                        dioxygenase; lethal leaf-spot 1"
                        /codon_start=1
                        /product="LLS1"
                        /protein_id="AAC49678.1"
                        /db_xref="PID:g1935912"
                        /db_xref="GI:1935912"
                        /translation="MRATIPALSLLVTPRLPSLAVPLAGRLREGGRSRTRLRVAAPT
                        SVPGEAAEQAEPSTSAPESGEKFSWRDHWYPVSLVEDLDPSRPTPFQLLNRDLVTMKE
                        PKSGEMVALLDRCPHRLAPLSEGRIDETGCLQCSYHGWSFDGSGACTKIPQAMPEGPE
                        ARAVRSPKACAIKPPTLVSQGLLFVWPDENGWEKAAAATKPPMLKEFEDPAFSTVTIQ
                        RDLFVGYDTLMENVSDPSHIEFAHHKVTGRDRARPLTEFMESSGAWGYSGANSGNPR
                        ITATFEAPCTALNKIEIDTKLPIFGQDKWVTWICSFNIPMAPGKTRSIVCSARNFQF
                        IMPQRDFVLACRTWLRKFGNSQPPEWFGNTEQEALPSTVLSKREMLDRYEQLSLKCSS
                        CKGAYNAFQNLQKVEMGATVVCCAAAGIPPDVQLRLLIGAALVSAAIAYAFHEL"

misc_feature            /note="encodes proline-cysteine motif suggested to be
                        important in coordinating the [4Fe-4S] iron-sulfur center"
                        3765..3853
                        join(3738..3764,3854..3904)
                        /gene="lls-1"

intron                  /gene="lls-1"
                        /number=1

exon                    3854..4089
                        /gene="lls-1"
                        /number=2

intron                  4090..4177
                        /gene="lls-1"
                        /number=2

exon                    4178..4304
                        /gene="lls-1"
                        /number=3

intron                  4305..5479
                        /gene="lls-1"
                        /number=3

exon                    5480..5616
                        /gene="lls-1"
                        /number=4

intron                  5617..5728
                        /gene="lls-1"
                        /number=4

exon                    5729..5888
                        /gene="lls-1"
                        /number=5

intron                  5889..6118
                        /gene="lls-1"
                        /number=5

exon                    6119..6397
                        /gene="lls-1"
                        /number=6

intron                  6398..6922
                        /gene="lls-1"
                        /number=6

exon                    6923..7129
                        /gene="lls-1"
                        /number=7

BASE COUNT              1869 a   1596 c   1543 g   2120 t
ORIGIN

Query Match            100.0%; Score 2822; DB 28; Length 7129;   1 others
```

```
Best Local Similarity 100.0%;  Conservative 0;  Pred. No. 0.00e+00;  Mismatches 0;  Indels 0;  Gaps 0;
Matches 2822;

Db    293  GCAACGCACACAGACAGGCACGCCACGGATCGTGCTCTTCGCGGGTCAGTAAACCTCACTCACACAGG  352
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1  gcaacgcacacagacaggcagcgatcgtcttcgcgggtcagtaaacctcactcacacagg   60

Db    353  CTATTCGTCTTAAGTTTTTTTTGTTCAACATCACATACTTGTGTTGCTTAATGTAACAAAA  412
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61  ctattcgtcttaagttttttttgttcaacatcacatacttgtgttgcttaatgtaacaaaa  120

Db    413  AAAATTCACACGCCCTCACAAACATTACAATATGAATTCAAAATAGAACTAACCAAACCTTG  472
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121  aaattcacacgcctcacaaacattacaatatgaattcaaaatagaactaaccaaaccttg  180

Db    473  GAGGACTTTGTACTGGCTAGAGAAGAACACCTACTCACTGCTTAACTGCTTACCCGAGACAG  532
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181  gaggacttgtactggctagagaagaacacctactcactgcttaactgcttacccgagacag  240

Db    533  AGGAAATACACGAGCAAACTCTTGTGGACTTGTTGCAAAATAGCAAGGAAAGGTATTAG  592
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241  aggaaatacacgagcaaactcttgtggacttgttgcaaaatagcaaggaaaggtattag  300

Db    593  TAATACGAAGCAATAAATTGTAGGAGCTGCAACGTATAACAATAACAATGATGAGTTGTAC  652
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301  taatacgaagcaataaattgtaggagctgcaacgtataacaataacaatgatgagttgtac  360

Db    653  CTTACATGTATGAAAATAAAAAAACTATATAGGTAAAGTGAACAAACATGCGTTATGTAAATC  712
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361  cttacatgtatgaaataaaaaactatataggtaaagtgaacaaacatgcgttatgtaaatc  420

Db    713  TAGCAGACTATTGGATTGAAAGAAATTCAATTACAAGGACAAAGAAGAAGACTGACGAGGGC  772
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421  tagcagactattggattgaaagaaattcaattacaaggacaaagaagaagactgacgagggc  480

Db    773  AGCAACACAAATAACTAAAATGTTCCAAAATGTGTCAGTATATGAAGGGCTCGAACGCACATGCAC  832
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481  agcaacacaataactaaatgttccaaaatggtcagatatgaagggctcgaacgcacatgcac  540

Db    833  GGCATGATAATGCTAGTTGGGGGCCGTTCCGTCGGGGCTTTAAAGTAGAGATAAGGAAAATCTGGATA  892
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541  ggcatgataatgctagttggggggccgttccgtcgggcgtttaaagataagggctcgaacagcac  600

Db    893  TGGACTAAATGAATGTCTAAATTTTTTGTTGTGAGACCTAGCCGCCATCGTAACTAGTAGAAGGT  952
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601  tggactaaatgatgtctaaatttttgttgtgagacctagccgcccctagagctaactagaaggt  660

Db    953  TAATTTTTGTTTCTATTTTTTTGTTGCACCGACGACACATTCTTTCTTTTGTCTCTAGTGTTT  1012
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661  taatttgttctctattttttgttgcaccgacgacacattcttttgttctaactagtagttt  720

Db   1013  ACATTTTAGTTACTACTCTCTTCCGTCCTCTTCGTGTCTAAAAAGTACTACTATATAAAAAATGTC  1072
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721  acatttagttactactctcttccgtctcaaaaagtactactatatctccattttttaaaatgtc  780

Db   1073  TTGCTTTTTTGAAGAGCACTACTTCTTTTTTTAAAATCTTGACCAACTACTATAAAAAAGTACTTCTGA  1132
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781  ttgcttttttgaagagcactactctctttttaaaatcttgaccaactataaaaagtacttctga  840

Db   1133  TACATGATAAGGTTTAATAAAAATATATAAAAATCTTATATTTTTAGTGAAGTCTTAGTGAAAACT  1192
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841  tacatgataggtttaataaaaatatataaaatcttatattttagtaagtctaagtgaaaact  900

Db   1193  TAAGAGCTTTGAATGTCGCACATAGTTGTTTTAAACAAGGTGTTTGTTCATGTTCGTTTCT  1252
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901  taagagctttgatgtcgcacatagttgttttaaacaaggtgtttgttcatgttcgttct  960

Db   1253  AATATAGTTGGATAGTAATTCCGATTCATTCGCCAGAGGTGTGGCCTGTGTGGATAATTTGGTTAG  1312
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961  aatatagtggatagtaattccgattcattcgccagaggtgtggcctgtgtggatattttggttag  1020

Db   1313  AGCATCTTCAAGAAAAACCGTAAATCAACTCCAAAAACGTTTGAGCCTCCCAACAGTCC  1372
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1021  agcatcttcaagagaaaccgtaaatcaactccaaaaacgtttgagcttccaacagtcc 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1373  ccccttccccttcccccccccATATATTACGCAAGTTGTTCCCAATCGGCCATIGCGG 1432

Qy   1081  ccccttcccccccccatattacgcgtcaaagcacattgttcccaatgctcctcgcatgctg 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1433  GTTCCCACGTGTATTTTCCTCGCGGCGGCAGTTCTGTTGGAGGAGGAGGCGGGACTTGGC 1492

Qy   1141  gttcccacgtgtattttcctcgcgcgcagttctgttggaggaggaaggcgggacgttggc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1493  ACTAGCGCTGGCGTGGAGATTATGGCCATCCGAATCACAGTTGTGTGCAGTCAAATGCTTTGT 1552

Qy   1201  actagcgctggctgagagattatggccatcgcaatcagtttgttggcagtcaaatgctttgt 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1553  TTTTTTTGGCCGCTCATGTGAGTGATCATTTTCTGTGAAAACTTAAATCAAATAATGAATGT 1612

Qy   1261  tttttttggccgctcatgtgagtatcatctgtgaaaactaaatcaataatgtgac 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1613  ATAATTCTTTAAGTCGTCACGATAGGAAGGAAGACTCCATCGTGTTCTCTGTGAAAACCTAAACCATGCAC 1672

Qy   1321  atattctttaagtcgtcacgatagggaggaagactccatcgttctaaacctaaaccatgcac 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1673  ACATATTCATCTTTCTTCCAAAACGCAAGTCTCGTGATATTTATATTTCTCGTGCCAGCTAGA 1732

Qy   1381  acatattcatctttctctccaaacgcaagtctcgtgatatttatattctcgtgccagctaga 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1733  TTAACTAGAAATTTAGATTCTTAAAAACTCTTTTAGAAAAAAAATATATCCAAACAGGA 1792

Qy   1441  ttactagaaaatttagattcttaaaaaactctttttagaaaaaaaatatataccaaacagga 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1793  CCATGGTTTAAACTATTACGGATAAAATAGCATGACTACCTTAGTAGTTTAAAATATATCAG 1852

Qy   1501  ccatggtttaaactattacggataaaatagcatgactaccttagtattataaaatgatatcag 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1853  TTGAAATATCTCGACTTATTTTATTGGTTAGACATGTTTAAATAATTATCA 1912

Qy   1561  ttgaaatatgtcgacttatttatttagttagtattattaggacatgtttaaataattatca 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1913  CAATTAAAACCAGATCTACATATAAACTATATTTGCTGTTGTCGTCAACTCGATCCGAAACTCACTT 1972

Qy   1621  catttaaaccagatctacatataaactatttgtcttgtcaactgatcgcaaactcactt 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1973  GCCTACCATCGGGATCGGATCGCGTATACAAGTGACACACTTTAAATGATTAAGCGCCGGA 2032

Qy   1681  gcctaccatcgggatcgcgctcgtatacaagtgacacacttttaaatgattaagccgcga 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2033  AAAATTATAAATGTACCATCCTCATTTGGCAAGTCAAAGAGAAAGCTTTACCATACAAATGA 2092

Qy   1741  aaaattataaatgtaccatccctcatttggcaagtctaaaggtacttttaccataaaatga 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  aactaaaattaaaattccaagtaataatataataataaaaaactgattttttcagttat 1860

Qy   1861  atattttagcaagctcgctaaatctgaatctagaaagttttttttgaaatgagatgatgct 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2153  ATAATTAGCAGCTGCTAAATCTGAATTTAGAAAGTTTTTTTGAAATGAGTGAGATGCT 2212

Qy   1921  cttataatggttacatatagggttgaggaggacggaagtagtagtagaaactggttgaagttttaaacaaaattc 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2273  GAATTTGATCTATTCAAACTTTGTAGCTACTCAGCGAGATGCGGAATTGCAAACATCGGGCG 2332

Qy   1981  gaatttgatctattcaactttgtagctactcagcaagatggaaattgcaaacatccggcg 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2333  GGGTGGAGTTCCGCCACGGCCCACGGGGTGGGTTCGTGTGTGTTCGTTCTCACCGCGGTCAATCTC 2392

Qy   2041  gggtggagttccgccacggcccacgggtgggttcgtgtcgttctcaccgccgggtcaatctc 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2393  CCCTCCGCGCGGCGAATTCGTCCCGGTCGGGACGGCGTAGCTGGCCCAATCCCAAAGCTC 2452

Qy   2101  ccctccgcgcgggcgaattcgtccggtggggacggcgtagctggcccaatcccaaagctc 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2453  CACCGACAAATGCCGCAAAGCGCCCATGCGTGTGGGTCGCGGGTACAAAITGCCTCCTTCCCCGCCC 2512

Qy   2161  caccgacaaatgccgcaaagcgccatgcgtggtcgcgtacaaattgcctcctttccccgccc 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2513  TTCCTCCCTTCCCTGCGGTGACGCACGCAACCACAACTGACGTCACCATGTGTGTACAAATGTATTC 2572

Qy   2221  ttcctcccttccctgcgtgacgcaccacaactgacgtcaccatcgtgtacaatgtattc 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2573  TCCCTAGCCGAACCGTATCACGTAGTTCTTAGGGTGGCGGTGCGTTCGCGGTTTACCGAAATTTT 2632

Qy   2281  tccctagccgaaccgtatcagttcgtcttagggttgtggcgttcgcggttaccggaaatttt 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2633  CGGGTTGGGTGAATTCAAGTTTTTTAAATTTCGGGTTTTGAGAAGTCAATAACTACCCGGAATTAC 2692

Qy   2341  cgggttgggtgaattcaagtttttttaaattcgggtttttgagaaatcaatacccggaattac 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2693  AACGGAATTTTTCAAATACCCGGAAATTTTCGGGTTCCGGGAAATTTCGGGTTCGGGTA 2752

Qy   2401  aacggatttttcaaatacccggaaattcgggttccgggttcgggta 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2753  TTCCCAAAACTACCCGAACTATTTGTTGTTGGGCTTCAGAAAAAAAACAATACACCCTATTAAATT 2812

Qy   2461  ttcccaaactacccgaactatttgttgttgggcttcataaaaaacaatacaccctattaaatt 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2813  AGTATAAAAAATATAGTTTGAATAAATAATAATGATGGACATGACATATTTAAAACACAACAACTAC 2872

Qy   2521  agtataaaaatatagtttgaataaataatgatataatacatgttataaaacacaacaactac 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2873  AATCCCAAGTTATGCACACTTACACATTACACATAATTAATAATAGATGTACAACTTAAATTATTATAAGC 2932

Qy   2581  aatcccaagttatgcacacttacacaataataatagatgtacaacttaaattattaaagc 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2933  ATGACATGAGTACATACACGACACATGAAAAGCCGGGGTAATTCGGGTATTCGGTTACCCGATTG 2992

Qy   2641  atgacatgagtacatacacatgaaaagccggggtaattcgggtattcgggttacccgattg 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2993  TGATACCCGAATTACCCGAAATAATAATTAATTAATTCGGGTTTGCTACCCGAAATTCCCAA 3052

Qy   2701  tgataccccgaattacccgaaataattaattcgggtttgctaccgaaatcccaa 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3053  ACAAAATTCGGGTTCGGGTATTCGGGTTCGGGTATTCCAGGTTGTTGGGGTTTCG 3112

Qy   2761  acaaaatcgggttcgggtattcgggttcgggttgttgggttttcg 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3113  GG 3114

Qy   2821  gg 2822
```

```
JOURNAL    Unpublished
REFERENCE  3 (bases 1 to 5364)
  AUTHORS  Kim,J.Y., Mahe,A., Tallercio,E.W., Chourey,P.S. and Prioul,J.L.
  TITLE    Direct Submission
  JOURNAL  Submitted (23-FEB-1998) Institut de Biotechnologie des Plantes,
           CNRS-ERS 569, Universite Paris-Sud (XI), Bat 630, Orsay 91405,
           France
FEATURES             Location/Qualifiers
     source          1..5364
                     /organism="Zea mays"
                     /strain="W22"
                     /db_xref="taxon:4577"
                     /chromosome="5"
                     /map="between gsy60e (bt2) and bnl771"
     mRNA            join(<1..208,821..829,1441..2306,2548..2706,3253..3509,
                     3636..3725,3818..>4003)
                     /product="cell wall invertase Incw1"
     CDS             join(1..208,821..829,1441..2306,2548..2706,3253..3509,
                     3636..3725,3818..4003)
                     /codon_start=1
                     /product="cell wall invertase Incw1"
                     /function="sucrose hydrolysis"
                     /note="beta-fructosidase"
                     /EC_number="3.2.1.26"
                     /db_xref="PID:g4105721"
                     /db_xref="GI:4105721"
                     /protein_id="AAD02511.1"
                     /translation="MGTRPGVVLAPWAVVLVLVLALRLAGASHVIHRSLEAEVPP
                     SVPASIVSPLLRTGVHFQPPMNWINDPNAPLIYKGWHLFYQYNPKGAVWGNIVAHS
                     VSRDLINWVALEAAIYPSISPSDKYGCWSGSATILEDGTPAILYTGIDRPDINYQOVL
                     ALPKDASDPLLREWKPEEVNPVATPPAAGINATQFRDPTTAWRHAGAGHWRMLVGSVR
                     GARGMALVYRSRDFRTWTKAKHPLHSAALTGMWECPDFFPVSGPGLQAGLDTSAPGTK
                     YVLKSSLDLIRYRYTIGSYDGDGDKDRYEPDDPAGDYRRLRYDVGNYYASKTFYDPVE
                     RRRVLLGWANESDSVPDDRAKGWAGIHAIPRKIWLDPTGKQLLQWPIHEVEKLRGKAV
                     SVDARLVKPGDHFEVTGIATQADVEVSFELEAGTSLLERAEAFDPAYDDAQKLCGV
                     KGADARGVGPFGLWVLASADLQERTAVFRVFRDGHGPKPKVLMCTDPTKSSLSPDLY
                     KPTLAGLVDADISSKRITLRSLIDRSVVESFGAGGKTCILSRVYPSIAVGKDAHLYVF
                     NNGEVDVTVSGLTAWEMKKPLMNGA"
     polyA_signal    4296..4301
BASE COUNT     1255 a   1352 c   1403 g   1354 t
ORIGIN
```

```
Query Match              9.7%;
Best Local Similarity   92.8%;    Score 275;  DB 28;   Length 5364;
Matches   298;  Conservative  0;  Mismatches  23;  Indels  0;  Gaps  0;

Db  5043 GTGCCTAGGGCGTGGGGCGTGGGGTTCGGGATTACCCGAAAAATTTCGGGCGTCGGGGTATTCGGGTGTTTTT 5102
         ||||| |||||||| ||||||||||||||||||||||||||||||||||||| ||||||| |||||||||||||
Qy  2304 gttcttagggcgtggggcgtggggttcgggattacccgaaaaatttcgggtgttgggtaattcaagttttt 2363
Db  5103 AAAAATTTTGGGTTTTTGGTTTTTTGAGAATCGCTCACCCCAAAATTACACAAGGGGTTTTTGCAATACCCGAAA 5162
         ||| | |||||| |||| ||||||||||||||  |||||||||||||||||| |||  |||||||||||||||||
Qy  2364 taaattcgggttttgagaatcaaatcaccccgaaattacacaacggatttttccaatacccggaa 2423
Db  5163 ATTCGGGTACCCAGAAATTCGGGTTCGGGTTATTCCCGACTACCCGAACTATTA 5222
         ||||||||||| |||||||||||||||||||||||||||  ||||||||||||||
Qy  2424 tttcgggtacccggaattcgggttcgggttcgggttattcccaaactaccgaactattg 2483
Db  5223 TGTCGGCCTTCATAAAAACACATCACCCTATTAAATTAGTATAAAATATAGTTTGAATA 5282
         ||| || |||||||||||| |||| ||||| |||||||| || |||||| ||||||||||
Qy  2484 tgttggcttcataaaaacaacataccacccctattaaatataaaaataatagttcgaata 2543
Db  5283 ATGATATACATGGACATATAAAAACACAACATCACACCCTATTAAATTAGTATAAAATAATAGTTTGAATA 5342
         |||||| ||||||||||| ||||| ||||||||  || |||||||||||||||||||||||| ||||||||||
Qy  2544 atgatacatggacataaaaacaacatctacacacaaacctattaaattagtataaaataatagtttgaata 2603
Db  5343 ACATAATTATAGATGCACAAA 5363
         ||||||||||||||| |||||
Qy  2604 acataattatagatgtacaaa 2624
```

```
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
NID         g2724471
VERSION     I66494.1  GI:2724471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheifflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1..7218
                     /organism="unknown"
BASE COUNT     1944 a   1491 c   1486 g   1929 t
ORIGIN
```

```
Query Match              2.9%;
Best Local Similarity   0.8%;    Score 83;  DB 25;  Length 7218;
Matches   3;  Conservative  219;  Mismatches  139;  Indels  0;  Gaps  0;

Db  1064 GATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1123
         ||| :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Cp   594 gatttcctatcttaaagcccgacggaaacgcccaaactagcatatcatgccgtgcat 535
Db  1124 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1183
         :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Cp   534 gcgttcgagcccctcatcatctgaccatttggaacattagttagttagttgtctgcccctc 475
Db  1184 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1243
         :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Cp   474 gtcagtcattctttgtccctgtaattctcttcaaatccaaatagtctgctgagatta 415
Db  1244 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1303
         :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Cp   414 cataacgcatgttgttcaacttaacctatatagtcttttatttcaatacatgtaaggtacta 355
Db  1304 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1363
         :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Cp   354 aagagcagactatcattgttatcacttgcagtccctacaaattatgcttgctattactaata 295
Db  1364 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1423
         :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Cp   294 cctttccttgctatttgcaacaacaagtccacacaacagttgtcgtgtgtattcctctgtct 235
Db  1424 Y 1424
         :
Cp   234 c 234
```

```
RESULT  4
LOCUS       MZESUT282     250 bp    DNA           PLN    08-MAY-1993
DEFINITION  Maize sucrose synthetase gene mutant (sh-5586 allele), containing
            transposon Tz86, 3' end of Tz86.
ACCESSION   M10175
NID         g168630
VERSION     M10175.1  GI:168630
KEYWORDS    sucrose synthase 3; transposon.
SEGMENT     2 of 2
SOURCE      Z. mays DNA.
  ORGANISM  Zea mays
            Eukaryota; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Poales; Poaceae; Zea.
REFERENCE   1 (bases 1 to 250)
  AUTHORS   Dellaporta,S.L., Chomet,P.S., Mottinger,J.P., Wood,J.A., Yu,S.M.
            and Hicks,J.B.
  TITLE     Endogenous transposable elements associated with virus infection in
            maize
  JOURNAL   Cold Spring Harb. Symp. Quant. Biol. 49, 321-328 (1984)
  MEDLINE   85153004
COMMENT     See segment 1.
```

```
RESULT  3
LOCUS       I66494        7218 bp    DNA           PAT    23-DEC-1997
```

```
FEATURES             Location/Qualifiers
     source          1..250
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
     repeat_region   231..240
                     /note="10 bp duplication"

BASE COUNT     69 a     57 c     54 g     70 t
ORIGIN    About 3.4 kb after segment 1.

Query Match     2.7%;    Score 75;    DB 27;    Length 250;
Best Local Similarity   77.4%;    Pred. No. 1.13e-29;
Matches  127;   Conservative    0;   Mismatches  34;   Indels   3;   Gaps   3;

Db      64 TCGGGTAGAACGGGC--TACCCGAACCCGAACCCGAATTTTTTGGGGTACCCGAATTTTCGG 122
           ||||||||||       |||  ||||||||||||||||||| ||||||||||||||||| |||
Cp    2476 tcgggtagttcgggaatacccgaacccgaacccgaaattcc-gggtacccgaaattcccgg 2418

Db     123 GTTTTGCTTAAATTGGGTACATATTCGGGTAITGTTTCTGAAAAACCCGAATTGTTAAAAG 182
           ||||| || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp    2417 gtattgaa-aaatccgttgtaattcgggtattgtttctgaaaaacccgaaatttaaaaaa 2359

Db     183 CCCAAAACTACCCGACCCGAAAAATTCGGGCCTRACCCGAATGCCCA 226
           | ||||| |||||||||||| |||||| ||||||||||||| ||||||
Cp    2358 cttgaattacccaacccgaaaattcgggtaaccgaacgcca 2315

RESULT     5
LOCUS          I6694        7218 bp     DNA        PAT        23-DEC-1997
DEFINITION   Sequence 14 from patent US 5670367.
ACCESSION    I6694
NID          g2724471
VERSION      I6694.1    GI:2724471
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 7218)
  AUTHORS    Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE      Recombinant fowlpox virus
  JOURNAL    Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1..7218
                     /organism="unknown"

BASE COUNT    1944 a    1491 c    1486 g    1929 t     368 others
ORIGIN

Query Match     2.1%;    Score 59;    DB 25;    Length 7218;
Best Local Similarity   1.6%;    Pred. No. 6.69e-19;
Matches   6;   Conservative   211;   Mismatches 158;   Indels    0;   Gaps    0;

Db    1426 YYYYYYYYYYGTACCA 1440
           ::::::::::||  |||
Qy    1420 tatattccgtgcca 1434

RESULT     6
LOCUS          AF012089    10772 bp    DNA        INV        05-AUG-1997
DEFINITION   Drosophila melanogaster cysteine proteinase-1 (CP1) gene, complete
             cds, and phenylalanyl tRNA synthetase gene, partial cds.
ACCESSION    AF012089
NID          g2305220
VERSION      AF012089.1    GI:2305220
KEYWORDS     .
SOURCE       fruit fly.
  ORGANISM   Drosophila melanogaster
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
             Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera;
             Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1  (bases 4546 to 4553)
  AUTHORS    Gray,Y.H., Tanaka,M.M. and Sved,J.A.
  TITLE      P-element-induced recombination in Drosophila melanogaster: hybrid
             element insertion
  JOURNAL    Genetics 144 (4), 1601-1610 (1996)
  MEDLINE    97132596
REFERENCE    2  (bases 1 to 10772)
  AUTHORS    Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
  TITLE      Structure of the cysteine proteinase (CP1) gene of Drosophila
             melanogaster and associated mutational effects
  JOURNAL    Unpublished
REFERENCE    3  (bases 1 to 10772)
  AUTHORS    Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
  TITLE      Direct Submission
  JOURNAL    Submitted (30-JUN-1997) School of Biological Sciences, University
             of Sydney, Biology A12, Sydney University, NSW 2006, Australia
FEATURES             Location/Qualifiers
     source          1..10772
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
     mRNA            join(872..1000,2310..2426,6476..6690,6751..7707)
                     /gene="CP1"
                     /product="cysteine protease"
     gene            872..7707
                     /gene="CP1"
     exon            872..1000
                     /gene="CP1"
                     /number=1
     intron          1001..2309
                     /gene="CP1"
                     /number=1
     exon            2310..2426
                     /gene="CP1"
                     /number=2
     CDS             join(2328..2426,6476..6690,6751..7462)
                     /gene="CP1"
                     /codon_start=1
                     /product="cysteine proteinase-1"
                     /protein_id="AAB65749.1"
                     /db_xref="PID:g2305221"
                     /db_xref="GI:2305221"
                     /translation="MRTAVLLPLLALLAVAQAVSFADVVMEEWHTFKLEHRKNYQDET
                     EERFRLKIFNENKHKIAKHNORFAEGKVSFKLAVNKYADLLHHEFRQLMNGFNYTLHK
                     QLRAADESFKGVTFISPAHVTLPKSVDWRTKGAVTAVKDQGHCGSCWAFSSTGALEGQ
                     HFRKSGVLVSLSEQNLVDCSTRYGNNGCNGGLMDNAFRYIKDNGGIDTEKSYPYEAID
                     DSCHFNKGTVGATDRGFTDIPQGDERKMAEAVATVGPVSVAIDASHESFQFYSEGVYN
                     EPQCDAQNLDHGVLVVGFGFDESGEDYWLYKNSWGTTWGDKGFIKMLRNKENQCGIAS
                     ASSYPLV"
     intron          2427..6475
                     /gene="CP1"
     misc_feature    4546..4553
                     /gene="CP1"
                     /note="insertion site of P[CaSpeR](50C)"
     exon            6476..6690
                     /gene="CP1"
```

(alignment block, left column)

```
Db     64 TCGGGTAGAACGGGC...
...
Db   1066 TYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1125
          |:||
Qy   1060 ttttgagcctccaacagtcccccttccccctcccccatattacgcgtcaagcattgttccc 1119

Db   1126 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1185
          ::|:::::||:::::::::::::::
Qy   1120 aatcgtcctctgcgcatgctgtccacgtgtatttcctcgcgcgacgcagttctgttgga 1179

Db   1186 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1245
          :::::::::::::::::::::::::::
Qy   1180 ggaggaaggcgggacgttggcactagcgctgctggagattatggccatcgcaatcagtt 1239

Db   1246 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1305
          :::::::::::::::::::::::::
Qy   1240 tgtggcagtcaaatgcttgttttttggccgctcatgtgagtatcattctgtgaaac 1299

Db   1306 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1365
          :::::::::::::::::::::::::
Qy   1300 tactcaaatcaatatgaatgtatattcttaagtcgtcacgataggaagactccatcgt 1359

Db   1366 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1425
          :::::::::::::::::::::::::
Qy   1360 tctaaaacctaaaccatgcacacatattcatcttttcccaaacgcaagtccgtgtgatatt 1419
```

Left column

```
Intron      /gene="CP1"
            /number=3
            6691..6750
exon        /gene="CP1"
            6751..7707
            /gene="CP1"
            /number=4
mRNA        /product="phenylalanyl tRNA synthetase"
            join(<8110..9300,9370..>9532)
            /gene="CP1"
CDS         /codon_start=1
            join(8110..9300,9370..>9532)
            /note="potential orf"
            /product="phenylalanyl tRNA synthetase"
            /protein_id="AAB65750.1"
            /db_xref="PID:g2305222"
            /db_xref="GI:2305222"
            /translation="MLITLRVQGARHWLKSTRCLASSAAPAKSPSSPPQLEVSGGSTYA
            TDGWTNVTPKILSYVGANKHLQTDHPLSIIRQRIVNFYGAYRNQRGNPLFSVYDQMN
            PVVTVQQNFDNLLIPADHVSRQKSDCYINQQHLLRAHTTAHQVELLSGGLDNFLVVG
            EVYRRDEIDSTHYPVFHQADAVRLVTHDKLFERNPGLELFEETWSGTLADRKLILFHP
            SSWTKNSPAYRRAVKLMEHEMKHVLVGLTKDLFGPRIKYRWVDTYPPFTQPSWELEI
            YFKDNWLEVLGGGIMRHEILQRSGVHQSIGFAFGVGLERLAMVLFDIPDIRLFWSNDS
            GFLSQFSEKDLINLPKYKPISHYPQCTNDLSFWLPQDIEVDAGFSPNDFYDLVRSVAG
            DMVEQISLVDKFKHPKTGKSSVCFRIVYRHMERTLTQAEVNEIHKQIASASVDSFNVQ
            IX"
```

```
BASE COUNT    2929 a    2357 c    2282 g    3046 t    158 others
ORIGIN
```

```
Query Match      1.9%;    Score 53;   DB 21;   Length 10772;
Best Local Similarity 21.2%;   Pred. No. 5.10e-15;
Matches  29; Conservative  71; Mismatches 35; Indels  2; Gaps  2;

Db  1694 WWYWAWTTMWKWMWTTMWKWAWMMKTYRTWWMWKMYWISRTTTTSAMMWMYTWSTWTKYWW 1753
         :!!:!!!!:!::!!:!!      !!:!:!:::::: : :!!::
Cp   883 taaaaatataagatttatatatttattaaacctatcatgtatcagaagtactttt-at 825

Db  1754 AYAWMKWWWTRTWARMAWASWARWKWTSAAAAYSAWRKMWKWWAYRAMKKTWMWAAWKW 1813
         !:!      !:!!::!!!  !:!!:!!!!!!!::!!!:
Cp   824 atagttggtcaagatttaaagatagtgctctcaa-aagcaagacatttaaaaat 766

Db  1814 RWKAAWWTWRWWYMTTT 1830
         : !!::!!::!! :!!!
Cp   765 gggagatagtacttt 749
```

```
RESULT     7
LOCUS      AF012089   10772 bp   DNA   INV   05-AUG-1997
DEFINITION Drosophila melanogaster cysteine proteinase-1 (CP1) gene, complete
           cds, and phenylalanyl tRNA synthetase gene, partial cds.
ACCESSION  AF012089
NID        g2305220
VERSION    AF012089.1  GI:2305220
KEYWORDS   .
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
           Eukaryotae; Metazoa; Arthropoda;
           Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera;
           Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1  (bases 4546 to 4553)
AUTHORS    Gray,Y.H., Tanaka,M.M. and Sved,J.A.
TITLE      P-element-induced recombination in Drosophila melanogaster: hybrid
           element insertion
JOURNAL    Genetics 144 (4), 1601-1610 (1996)
MEDLINE    97132596
REFERENCE  2  (bases 1 to 10772)
AUTHORS    Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE      Structure of the cysteine proteinase (CP1) gene of Drosophila
           melanogaster and associated mutational effects
JOURNAL    Unpublished
REFERENCE  3  (bases 1 to 10772)
AUTHORS    Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE      Direct Submission
JOURNAL    Submitted (30-JUN-1997) School of Biological Sciences, University
```

Right column

```
           of Sydney, Biology A12, Sydney University, NSW 2006, Australia
FEATURES            Location/Qualifiers
source      1..10772
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"
mRNA        join(872..1000,2310..2426,6476..6690,6751..7707)
            /product="cysteine protease"
            /gene="CP1"
gene        872..1000
            /gene="CP1"
            /number=1
exon        872..1000
            /gene="CP1"
            /number=1
intron      1001..2309
            /gene="CP1"
            /number=1
exon        2310..2426
            /gene="CP1"
            /number=2
CDS         join(2328..2426,6476..6690,6751..7462)
            /codon_start=1
            /note="potential orf"
            /product="cysteine proteinase-1"
            /protein_id="AAB65749.1"
            /db_xref="PID:g2305221"
            /db_xref="GI:2305221"
            /translation="MRTAVLLPLLALLAVAQAYSFADVVMEEWHTFKLEHRNYQDET
            EERFRLKIFENKHKIAKHNQRFAEGKVSFKLAVNKYADLLHHEFRQLMNGFNYTLHK
            QLRAADESFKGVTFISPAHVTLPPKSVGDWRTKGAVTAVKDQGHCGSCWAFSSTGALEGQ
            HFRKSGVLVSLSEQNLVDCSTKYGNNGCGLMDNAFRYIKDNGGIDTEKSYPYEAID
            DSCHFNKGTVGATDRGFTDIPQGDEKKMAEAVATVGPVSVAIDASHESFQYSEGVYN
            EPQCDAQNLDHGVLVVGFGTDESGEDYWLVKNSWGTTWGDKGFIKMLRNKENOCGIAS
            ASYPLV"
misc_feature 4546..4553
            /gene="CP1"
            /note="insertion site of P[CaSpeR](5OC)"
            /citation=[1]
intron      2427..6475
            /gene="CP1"
            /number=2
exon        6476..6690
            /gene="CP1"
            /number=3
intron      6691..6750
            /gene="CP1"
            /number=3
exon        6751..7707
            /gene="CP1"
            /number=4
mRNA        join(<8110..9300,9370..>9532)
            /product="phenylalanyl tRNA synthetase"
            /gene="CP1"
CDS         /codon_start=1
            join(8110..9300,9370..>9532)
            /note="potential orf"
            /product="phenylalanyl tRNA synthetase"
            /protein_id="AAB65750.1"
            /db_xref="PID:g2305222"
            /db_xref="GI:2305222"
            /translation="MLITLRVQGARHWLKSTRCLASSAAPAKSPSSPPQLEVSGGSTYA
            TDGWTNVTPKILSYVGANKHLQTDHPLSIIRQRIVNFYGAYRNQRGNPLFSVYDQMN
            PVVTVQQNFDNLLIPADHVSRQKSDCYINQQHLLRAHTTAHQVELLSGGLDNFLVVG
            EVYRRDEIDSTHYPVFHQADAVRLVTHDKLFERNPGLELFEETWSGTLADRKLILFHP
            SSWTKNSPAYRRAVKLMEHEMKHVLVGLTKDLFGPRIKYRWVDTYPPFTQPSWELEI
            YFKDNWLEVLGGGIMRHEILQRSGVHQSIGFAFGVGLERLAMVLFDIPDIRLFWSNDS
            GFLSQFSEKDLINLPKYKPISHYPQCTNDLSFWLPQDIEVDAGFSPNDFYDLVRSVAG
            DMVEQISLVDKFKHPKTGKSSVCFRIVYRHMERTLTQAEVNEIHKQIASASVDSFNVQ
            IX"
```

```
BASE COUNT    2929 a    2357 c    2282 g    3046 t    158 others
ORIGIN
```

```
Query Match      1.7%;    Score 49;   DB 21;   Length 10772;
Best Local Similarity 19.0%;   Pred. No. 1.71e-12;
Matches  26; Conservative  72; Mismatches 37; Indels  2; Gaps  2;

Db  1689 AAAKRWWTWWAWTTTMWKWMWTTMWKWAWMMKTYRTWWMWKMYWISRTTTTSAMMWMYTWSTW 1748
         |||  :::::::  |    !||!:!!::! : :!|::::::! : :||:!:: !!:!: ! |:
```

```
Qy    749 aaaagtactatatccccatttttttaaaatgtcttgctt-tttgaagcactatctttta 807
Db   1749 TKYVWWAYAVWMKWMWWTRTWARMAWASWARWRKWKTSAAAAYSAWRKWWKWWWAYRAMKKTWMW 1808
Qy    808 aaatcttgaccaacctatataaaa-gtacctcctgacatgataggtttaataaaatat 866
Db   1809 AAWKRWRKAAWWTWRWW 1825
Qy    867 aaaatcttatattta 883
```

RESULT 8

```
LOCUS       MZESUT281    250 bp    DNA                PLN    08-MAY-1993
DEFINITION  Maize sucrose synthetase gene mutant (Sh-5586 allele) containing
            transposon Tz86, 5' end of Tz86.
ACCESSION   M10174
NID         g168629
VERSION     M10174.1  GI:168629
KEYWORDS    sucrose synthase 3; transposon.
SEGMENT     1 of 2
SOURCE      Zea mays DNA.
  ORGANISM  Zea mays
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Poales; Poaceae; zea.
REFERENCE   1  (bases 1 to 250)
  AUTHORS   Dellaporta,S.L., Chomet,P.S., Mottinger,J.P., Wood,J.A., Yu,S.M.
            and Hicks,J.B.
  TITLE     Endogenous transposable elements associated with virus infection in
            maize
  JOURNAL   Cold Spring Harb. Symp. Quant. Biol. 49, 321-328 (1984)
  MEDLINE   85153004
COMMENT     The transposon Tz86 generates a 10 bp duplication of target
            sequence upon integration. This sucrose synthetase allele (with
            the Tz86 insertion) been named sh-5586. [1] estimates Tz86 to be
            3.6 kb in length.
FEATURES             Location/Qualifiers
     source          1..250
                     /organism="zea mays"
                     /db_xref="taxon:4577"
     CDS             <1..41
                     /note="sucrose synthetase, (putative); putative"
                     /codon_start=3
                     /number=15
                     /protein_id="AAA35516.1"
                     /db_xref="PID:g168632"
                     /db_xref="GI:168632"
                     /translation="ASRYTGSSTPRG"
     repeat_region   34..43
                     /note="10 bp duplication"
BASE COUNT     84 a     74 c     49 g     43 t
ORIGIN      1 bp upstream of HindIII site.
```

```
Query Match     1.68;   Score 45;   DB 27;   Length 250;
Best Local Similarity   69.8%;   Conservative   0;   Mismatches 40;   Indels 2;   Gaps 2;
Matches 97;

Db     78 AAAACCCGAAAAACCCGAGCCCGAACCCGTATAGCCCCAATACCCGAPTACCCGAAAATATTTA 137
Cp   2812 aaacctggaataccogaaccogaacccgaaatccgaattctgttttgggaat 2753

Db    138 TGCGGGCAGTAAACCTTAAAACCCGATATATTATTTCGGGCTTAATCGGTAGCAACACCGG 197
Cp   2752 ttcgggtagcaacttgcaaaaccgaaattattccgg-taattcgggtatcacaatcgg 2694

Db    198 -TACCCGAAATACCCAAAT 215
Cp   2693 gtaccgaataccgaat 2675
```

RESULT 9

```
LOCUS       ZMCI31AC3    5057 bp    DNA                PLN    24-FEB-1999
DEFINITION  z.mays cyp71c3 gene.
```

```
ACCESSION   Y11403
NID         g1870202
VERSION     Y11403.1  GI:1870202
KEYWORDS    CYP71C3 gene; Cytochrome P450.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
            Poaceae; zea.
REFERENCE   1  (bases 1 to 5057)
  AUTHORS   Gierl.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1997) A. Gierl, TU-Muenchen, Genetic,
            Lichtenbergstrasse 4, 85747 Garching, FRG
REFERENCE   2  (bases 1 to 5057)
  AUTHORS   Frey,M., Chomet,P., Glawischnig,E., Stettner,C., Grun,S.,
            Winklmair,A., Eisenreich,W., Bacher,A., Meeley,R.B., Briggs,S.P.,
            Simcox,K. and Gierl.A.
  TITLE     Analysis of a chemical plant defense mechanism in grasses
  JOURNAL   Science 277 (5326), 696-699 (1997)
  MEDLINE   97382314
FEATURES             Location/Qualifiers
     source          1..5057
                     /organism="zea mays"
                     /cultivar="ci31a"
                     /db_xref="taxon:4577"
                     /chromosome="4"
                     /dev_stage="seedling"
     exon            2171..2736
                     /number=1
     gene            2203..4029
                     /gene="cyp71c3"
     CDS             join(2203..2736,2833..3273,3397..4029)
                     /gene="cyp71c3"
                     /codon_start=1
                     /product="cytochrome p450"
                     /protein_id="CAA72207.1"
                     /db_xref="PID:e304532"
                     /db_xref="PID:g1870203"
                     /db_xref="GI:1870203"
                     /translation="MALQAAYEYLQQAVGHGAWSSTQTLTLLLLIAVPTVLLLLASLAK
                     STSSSGRGKPLPPSPPGTIPIVGHLHHIGPOTHISLQELVAKYGHGFLFLRAGAVP
                     TLIVSSPSAAEAVMRIHDHICASRPWSMASHILRYNTCDVAFSPLGEYWQQTRKLMNT
                     HLLSNKKVSYPRHGREEEVCLLVDNLREAAKSPSTAVDMSEVLAAYTNDVVSRSVLG
                     STHRKGRNTLFREMTMTINVDLLVGENLEYYIPRWPLLTDLLFRLVCWKVTRHLKRVDA
                     LLEEVIHEHVEMRKLSGDKEKESDDIDIFLSRYEEYGFDNVKSLLMNVFEAAIET
                     SYLVLESAMAELMNHRVMKLQAEVRAYGAEKKLDMIREDDLSSLPYLKASMKEALR
                     LHPPGPLLLPHYSTADCOIDGYHIPANPRVLVNGWAIGRDPAVWKEPEEFMPERFMRD
                     GWDKSNSYSGQDFRILPFGSGRRICGANFGLATMIMLANLMYHFDWEVPNEKEDGC
                     WKVSMDEKFGLMLRRNELLIVPRASS"
     intron          2737..2832
                     /gene="cyp71c3"
                     /number=1
     exon            2833..3273
                     /gene="cyp71c3"
                     /number=2
     intron          3274..3396
                     /gene="cyp71c3"
                     /number=2
     exon            3397..>4029
                     /gene="cyp71c3"
                     /number=3
BASE COUNT   1263 a   1228 c   1270 g   1296 t
ORIGIN
```

```
Query Match     1.5%;   Score 43;   DB 27;   Length 5057;
Best Local Similarity   65.7%;   Pred. No. 8.02e-09;
Matches 90;   Conservative   0;   Mismatches 47;   Indels 0;   Gaps 0;

Db    925 GATTCCTTCTATCGCGACGACTGACCTATTCTGTTTTTAATGTTGATTGACGATTTC 984
Cp   1353 gagtcttcctatccgtcgacgacttaagaaatatacattcatattgattagatagttttc 1294
```

```
Db    985  ACTAAAGCGGATTCTCACGGAGAGTGGGCGCTGAAAAGCTAAGCGTTGGGCAGTTCCGCAACAGC  1044
             ||  |||||||||||||||||||  |||||||||||
Cp   1293  acagaaatgtatcctccacatgagcggcgccaaaaacaaacaaagcatttgactgccacaaactga  1234

Db   1045  TTCTCGGTGGCCAGAATC 1061
             |||||||||||||
Cp   1233  ttgcgatggccataatc  1217
```

RESULT   10

```
LOCUS       ZMA132240    6831 bp    DNA        PLN     15-JAN-1999
DEFINITION  Zea mays eIF-5 gene, exons 1-2.
ACCESSION   AJ132240
NID         g4160401
VERSION     AJ132240.1  GI:4160401
KEYWORDS    eIF-5 gene; eukaryotic translation initiation factor 5.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
            Poaceae; Zea.
REFERENCE   1  (bases 1 to 6831)
  AUTHORS   Ignacio,L.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1999) Ignacio L.R., Genetica Molecular, Instituto
            de Biologia Molecular de Barcelona, CID, CSIC, C/ Jordi Girona
            Salgado 18-26, 08034 Barcelona, ESPANA
REFERENCE   2  (bases 1 to 6831)
  AUTHORS   Ignacio,L.R. and Pere,P.R.
  JOURNAL   Unpublished
COMMENT     Exist repetitive fragments to 1 from 1556 and  to 5357 from 5752.
FEATURES             Location/Qualifiers
     source          1..6831
                     /organism="zea mays"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /chromosome="4L"
     promoter        1556..1901
                     /gene="eIF-5"
     gene            1556..5162
                     /gene="eIF-5"
     TATA_signal     1645..1651
                     /gene="eIF-5"
     TATA_signal     1665..1670
                     /gene="eIF-5"
     mRNA            1822..1828
                     /gene="eIF-5"
     exon            1902..1949
                     /gene="eIF-5"
                     /number=1
     intron          1950..3247
                     /gene="eIF-5"
                     /number=1
     mRNA            join(1902..1949,3248..5162)
                     /gene="eIF-5"
     exon            3248..5162
                     /gene="eIF-5"
                     /number=2
     CDS             3545..4900
                     /gene="eIF-5"
                     /codon_start=1
                     /product="eukaryotic translation initiation factor 5"
                     /protein_id="CAA10616.1"
                     /db_xref="PID:e1370838"
                     /db_xref="PID:g4160402"
                     /db_xref="GI:4160402"
                     /translation="MALQNIGASNRDDAFYRYKMPRMITKIEGRGNGIKINVANMAVDI
                     AKALARPASYTTKYFGCELGAQOSKFDEKTGISLVNGAHDTAKLAGLLEVFIRKTVQCY
                     GCGNPETEIIISKTQMISLKCAACGFVSDVDMRDKLTTFILKNPPEQKKGGKDKKAMR
                     RAEKERLKEGEAADEEQKLKKDAKKGSKDSTAKGLKKKRATTATGSDEDHSSSPTRS
                     HDCDKAAADDDDDDVQWQTDTSIEAAKQRMOEQLSAATAEMVMLSTEETEKKMKQPTH
                     KDGTNGSAKEIPNDKPAVTKPSPYEELICDIKASLGSAPTPSQLKAVLASSTLPPQD
                     VMNALLEALFGGVKGGFIKEVVKNKKYLAVAVPDEGAQOTLLVQAIEAFGKKCNPEALK
                     EVPVVLKALYDGDILDEETIVDWYNDAVAAGKDSQVVKNAKPFVEWLQSAESDEEGGD
                     E"
     polyA_signal    4951..4955
                     /gene="eIF-5"
     terminator      5138..5142
                     /gene="eIF-5"
BASE COUNT     1867 a   1368 c   1490 g   2106 t
ORIGIN
```

```
Query Match           1.5%;   Score 43;   DB 27;   Length 6831;
Best Local Similarity 60.7%;  Pred. No. 8.02e-09;
Matches   179;  Conservative   0;   Mismatches 112;   Indels   4;   Gaps   3;

Db    433  AGCTTCTGGGCCACCAAAAGCTGCTGCGGACTGCCAAAACGCTCAGCGTTTTCAGCCAGCTTC  492
           || |||| ||||||| || ||| ||| ||  |||||||||| ||||| ||||| ||||||
Qy   1216  agattatggccatcgcaatcagcttgtcggcagtcagtcaaatgcttgttttttggcggtca  1275

Db    493  TATAAAATTCGTTAGGAGGAACAAAAACCATCTAAAATCAACATAAAACACATAAATGGGTTGAGT  552
           ||||| ||||| ||| || ||||||| |||| ||||| |  |||| || ||| ||| |||| |||
Qy   1276  tgtgagtcattctcttgtgaaaacatatctaa-tcaaatatgaatgtatatattcttaagt  1334

Db    553  CGTTGTAATAGTAGGAAAATCCCGTCATTTTCTAGATCATGAGCAGCTATGAACAACTTTATCT  612
           || | | |||| ||| ||| ||| |||| ||  ||||  ||| || || |||  |||||| |||
Qy   1335  cgtcacgatagagaagactccatcgtt--ctaaaacctaaacaccatgcaacacatatccatct  1392

Db    613  TCCTCCACACGTAAATGGTAAAATGATACTACAAATTTTTCCCACAACCACAGATTCTCCCACAG  672
           ||| |||  || |||  ||||| || | |  | || | | | | |||  |||  ||||| |||
Qy   1393  tctctccaaacgcaagtctcgtgatatttatctcgtgccagctagaatattcctagaaat  1452

Db    673  CCAGGATTTTCAGAAAAGCTGACCAGAAAAAAGCTGA-ACCAAACAGCAGCCCCATAGT  726
           |||| ||||| || ||| | |  ||||  ||||||| | | |||||| || ||||| ||
Qy   1453  ttagattcttaaaaaaattacccaacaggaccatggt  1507
```

RESULT   11

```
LOCUS       AR024229    965 bp    DNA        PAT     04-DEC-1998
DEFINITION  Sequence 22 from patent US 5795961.
ACCESSION   AR024229
NID         g3977523
VERSION     AR024229.1   GI:3977523
KEYWORDS    
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 965)
  AUTHORS   Wallace,T.Paul, Harris,W.J., Carr,F.J., Old,I.J., Welt,S. and
            Kitamura,K.
  TITLE     Recombinant human anti-Lewis b antibodies
  JOURNAL   Patent: US 5795961-A 22 18-AUG-1998;
FEATURES             Location/Qualifiers
     source          1..965
                     /organism="unknown"
ORIGIN
```

```
Query Match           1.4%;   Score 39;   DB 25;   Length 965;
Best Local Similarity 16.8%;  Pred. No. 1.81e-06;
Matches    18;  Conservative   56;   Mismatches 32;   Indels   1;   Gaps   1;

Db    843  RSSVTAADIAVYYCVRGRSYDSDGGDIYWGGTIVTVSSHTVKDMTSSSSASVGDRVTICRS  902
           :::: || |:||:::|:|::::|:|::| :::::::::|:|:::::::|::::::
Qy   2714  accgaaataatttcgggttttgcaagttgctaccgaaatc-ccaaacaaaattcggg  2772

Db    903  STTHGNGNTIYWYKGKAKYRVSNRSGVSRSGSGSGTDYTTSSDATIY  949
           |:|:|::::::| :::::::::::: ::|:::: :::::|:|: :||:
Qy   2773  tttcgggtattccgggttcggggtttccaggtttcgggttcc  2819
```

RESULT   12

```
LOCUS       MVU87256    1056 bp    DNA        MAM     02-JAN-1999
DEFINITION  Mustela vison GT dinucleotide repeat, chromosome 1q.
ACCESSION   U87256
```

NID   g4099442
VERSION   U87256.1 GI:4099442
KEYWORDS   .
SOURCE   American mink.
  ORGANISM   Mustela vison
              Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
              Carnivora; Fissipedia; Mustelidae; Mustela.
REFERENCE   1 (bases 1 to 1056)
  AUTHORS   Brusgaard,K., Shukri,N.M., Malchenko,S., Koroleva,I. and Lohi,O.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-JAN-1997) Breeding and Genetics, Danish Institute of
              Animal Science, Blichersalle K25, Tjele 8830, DK
FEATURES   Location/Qualifiers
     source   1..1056
                     /organism="Mustela vison"
                     /db_xref="taxon:9667"
                     /chromosome="1"
                     /map="1q"
     primer_bind   98..119
                     /note="primers: 1167F: agcccctgcatctcacttctt, 1167R:
                     gaggatcttaccgctgttgag"
                     /standard_name="1167F"
     primer_bind   complement(300..320)
                     /standard_name="1167R"
BASE COUNT   211 a   221 c   210 g   225 t   189 others
ORIGIN

Query Match   1.3%;   Score 38;   DB 23;   Length 1056;
Best Local Similarity 17.6%;   Pred. No. 6.78e-06;
Matches   33; Conservative 87; Mismatches 64; Indels 3; Gaps 3;

Db   572 CCCWSCRHDVBMWSKWCWGKKSCCKSTGKGKDMSGCAYCGKRMRCSRYMMRMCVGSGGTW 631
           |||·|·|||:|·||:||·||||·||·|·||· ||·:|· ||·:||||· |||·:·:|||·
Qy   1070 cccaacagtccccctccccctccccacatatt-acgcgtcaagcaattgttcccaatcgtcct 1128

Db   632 ARCCCCDKKSGSGSGDKSHCKSKRKYKWMDRYHBCKSMCAMVRGKWSAGTGCMARGWRTWDS 691
           ·|·:|·:||||·|·||·||· ||·|·|· ||:|·||||·||·||·|·|·|·||·|·||·|·|
Qy   1129 ctgcgcatgctggttcccacgtgtattttcctcgcgcgcagttctgttggaggaggaagg 1188

Db   692 SCTSTKGARRRKGMYARRKMYGTGAAWASAT-AYSTGAGTTBWRVTSWKMBKYKKTGMMH 750
           :·| ·|||||·||·||·||·|||·|·||:|· |··|· |||·||· ||· |·||·:|||·
Qy   1189 cgggacgttggcactagccgtggctggagattatggccatcgcaatcag-tttgtggcag 1247

Db   751 HCAWKYG 757
           :||:·|·|
Qy   1248 tcaaatg 1254

RESULT 13
LOCUS   MZECPN60B   7905 bp   DNA   PLN   14-JUL-1993
DEFINITION Corn nuclear-encoded mitochondrial chaperonin 60 (cpn60II) gene,
            complete cds.
ACCESSION   L21008
NID   g309558
VERSION   L21008.1 GI:309558
KEYWORDS   chaperonin 60; nuclear-encoded mitochondrial protein.
SOURCE   zea mays (strain B73) (library: Dashii of P.S.Close) DNA.
  ORGANISM Zea mays
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Poales; Poaceae; Zea.
REFERENCE   1 (bases 1 to 7905)
  AUTHORS   Close,P.S.
  TITLE   Molecular characterization of two nuclear genes encoding zea mays
            mitochondrial chaperonin 60
  JOURNAL   Unpublished (1993)
FEATURES   Location/Qualifiers
     source   1..7905
                     /organism="zea mays"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /tissue_lib="Dashii of P.S.Close"
                     /map="1L"

     5'UTR   1711..1913
                     /gene="cpn60II"
     exon   1711..1982
                     /gene="cpn60II"
                     /label=exon1
     gene   1711..7905
                     /gene="cpn60II"
     CDS   join(1914..1982,2409..2540,2807..2966,4035..4210,
                     4299..4394,4698..4802,4903..4989,5080..5169,5231..5311,
                     5397..5465,5645..5722,5816..5911,6049..6148,6309..6412,
                     6492..6597,6700..6778,7024..7126)
                     /gene="cpn60II"
                     /codon_start=1
                     /product="chaperonin 60"
                     /protein_id="AAA33451.1"
                     /db_xref="PID:g309559"
                     /db_xref="GI:309559"
                     /translation="MYRAAASLASKAROAGSSSAARQVGSRLAWSRNYAAKDIKFGVE
                     ARALMLRGVEELLADAVKVTMGPKGRNVVIEQSFGAPKVTKDGVTVAKSIEFRDRVKNV
                     GASLVKQVANAINDIACDGTTCATVLAKAIFTEGCKSVAAGNNAMDLRRGISMAVDAV
                     VTNLKGMARMISTSEEIAQVGTISANGEREIGELIARAMEKVGKEGVITIADGNTLYN
                     ELEVVEGMKLDRGYISPYFITNSKAQKCELEDPLILIHDKKVTNMHAVVKVLEMALKK
                     ORPLLIVAEDVESEALGTLIINKLRAGIKVCAVKAPGFGENRKANCLDLAILTGGEVI
                     TEELGMNLENVEPHMLGSSCKKVTVSKDTVILDGAGDKKSIEERADQIRSAVENSISD
                     YDKEKLQERLAKLSGGVAVLKIGGASEAEVGEKKDRVTDALNATKAAVEEGIVPGGGV
                     ALLYASKELDKLQTANFDQKIGVQIIQNALKTPVHTIASNAGVEGAVVVGKLLEQGNT
                     DLGYDAAKDEVVDMVKAGIIDPLKVIRTALVDAASVSSLMTTTESIIVEIPKEEAPAP
                     AMGGMGGMDY"

     exon   2409..2540
                     /gene="cpn60II"
                     /label=exon2
     exon   2807..2966
                     /gene="cpn60II"
                     /label=exon3
     exon   4035..4210
                     /gene="cpn60II"
                     /label=exon4
     exon   4299..4394
                     /gene="cpn60II"
                     /label=exon5
     exon   4698..4802
                     /gene="cpn60II"
                     /label=exon6
     exon   4903..4989
                     /gene="cpn60II"
                     /label=exon7
     exon   5080..5169
                     /gene="cpn60II"
                     /label=exon8
     exon   5231..5311
                     /gene="cpn60II"
                     /label=exon9
     exon   5397..5465
                     /gene="cpn60II"
                     /label=exon10
     exon   5645..5722
                     /gene="cpn60II"
                     /label=exon11
     exon   5816..5911
                     /gene="cpn60II"
                     /label=exon12
     exon   6049..6148
                     /gene="cpn60II"
                     /label=exon13
     exon   6309..6412
                     /gene="cpn60II"
                     /label=exon14
     exon   6492..6597
                     /gene="cpn60II"
                     /label=exon15

```
exon             6700..6778
                 /gene="cpn60II"
                 /label=exon16
exon             7024..7383
                 /gene="cpn60II"
                 /label=exon17
3'UTR            7124..7383
                 /gene="cpn60II"
polyA_signal     7249..7254
                 /gene="cpn60II"
polyA_site       7383
                 /gene="cpn60II"
BASE_COUNT       2012 a   1516 c   1802 g   2575 t
ORIGIN
```

```
RESULT          14
Query Match
Best Local Similarity   65.9%;
Matches 108;  Conservative  0;  Mismatches 53;  Indels  3;  Gaps  3;

Score 37;  DB 27;  Length 7905;
Pred. No. 2.51e-05;

Db   337 TCGGGTAGAACGGGC-TACCCATACCCGCAATCGGAATTTTCTAGATACCCAAAATTTCAGG 395
         |||||||| | ||| ||| |||||| ||||||| |||||| |||||||||| ||| ||
Cp  2476 tcgggtagttcgggaataccccgaaccggaatttttctagataccccgaaattttcggg 2417

Db   396 TTTAGTTTAGATCGAGTACATTTTCGAGTATATTATTCTAAAAGCTCGAATTGTGAAAAGC 455
         |-|||| |||| || || ||| |||||||||| | ||||| | |||||||||| |||
Cp  2416 tattgaaa-atccgttgtaatttcgggtattgattctcaaaacccgaaattaaaaaac 2358

Db   456 CCAATCTTAGCCGACCCAAAAAAATACGGGCTAAATCAAAATGCCCA 499
         | ||  |||| |||||| ||||| |||| | |||||| ||| ||||
Cp  2357 ttgaattacccaaccgaaaatcggg-taaccgaacgcca 2315
```

```
LOCUS       AC005369    74371 bp    DNA    PRI    01-AUG-1998
DEFINITION  Homo sapiens chromosome 5, BAC clone 11j93 (LBNL H175), complete
            sequence.
ACCESSION   AC005369
NID         g3367505
VERSION     AC005369.1  GI:3367505
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 74371)
  AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
  TITLE     Sequencing of human chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 74371)
  AUTHORS   Ricke,D.O.
  TITLE     Large Scale Sequence Analysis and Annotation with the Sequence
            Comparison Analysis (SCAN) System
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 74371)
  AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
            Institute, Lawrence Berkeley National Laboratory, MS 74-157,
            Berkeley, CA 94720, U.S.A.
COMMENT     Sequence submitted by:
            DOE Joint Genome Institute.
FEATURES             Location/Qualifiers
     source          1..74371
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="5q"
                     /clone="11j93"
                     /chromosome="5"
                     /note="LBNL H175"
```

```
     repeat_region   893..1030
                     /rpt_family="Alu"
     repeat_region   2295..2438
                     /rpt_family="Alu"
     repeat_region   2818..2859
                     /note="(GT)21"
                     /rpt_type=tandem
     misc_feature    join(3246..3410,3721..3828)
                     /note="65% & 69% protein identity GenPept:U22377"
                     /standard_name="RLF"
     repeat_region   3431..3724
                     /rpt_family="Alu"
     repeat_region   3707..3728
                     /note="(A)22"
                     /rpt_type=tandem
                     /rpt_unit=A
     repeat_region   4366..4661
                     /rpt_family="Alu"
     repeat_region   5327..5602
                     /rpt_family="Alu"
     repeat_region   6586..6956
                     /rpt_family="L1"
     repeat_region   6647..6684
                     /note="(CA)19"
                     /rpt_unit=CA
                     /rpt_type=tandem
     repeat_region   7113..7373
                     /rpt_family="Alu"
                     /rpt_unit=CA
     repeat_region   complement(7830..8185)
                     /standard_name="possible repeat"
     repeat_region   8258..8503
                     /rpt_family="Alu"
     repeat_region   9070..9387
                     /rpt_family="Alu"
                     /rpt_type=tandem
                     /rpt_unit=A
     repeat_region   complement(9740..9845)
                     /rpt_family="Alu"
     repeat_region   complement(10440..11015)
                     /rpt_family="MER42"
     repeat_region   11950..12250
                     /rpt_family="Alu"
     repeat_region   12057..12085
                     /note="(A)29"
                     /rpt_type=tandem
                     /rpt_unit=A
     repeat_region   12365..12645
                     /rpt_family="Alu"
     repeat_region   13727..13750
                     /note="(AC)12"
                     /rpt_type=tandem
                     /rpt_unit=AC
     repeat_region   13783..14024
                     /rpt_family="L1"
     repeat_region   14175..14470
                     /rpt_family="Alu"
     repeat_region   complement(14906..15259)
                     /rpt_family="Alu"
     repeat_region   15300..15613
                     /standard_name="possible repeat"
     repeat_region   16671..16690
                     /note="(T)20"
                     /rpt_type=tandem
                     /rpt_unit=T
     repeat_region   complement(16675..16977)
                     /rpt_family="Alu"
     repeat_region   complement(16993..17085)
                     /rpt_family="Alu"
     repeat_region   complement(17678..18276)
                     /rpt_family="MER42"
     repeat_region   19305..19583
                     /rpt_family="Alu"
     repeat_region   19914..19945
                     /note="(GTT)8"
                     /rpt_family="Alu"
```

```
     repeat_region
                     /rpt_type=tandem
                     /rpt_unit=GTTT
                     complement(19943..20222)
     repeat_region   complement(21202..21496)
                     /rpt_family="Alu"
     repeat_region   21736..22035
                     /rpt_family="Alu"
     repeat_region   22017..22038
                     /note="(A)22"
     repeat_region   2231..22457
                     /rpt_type=tandem
                     /rpt_unit=A
     repeat_region   22231..22457
                     /rpt_family="Alu"
     repeat_region   complement(22638..22981)
                     /standard_name="possible repeat"
     repeat_region   23022..23326
                     /rpt_family="Alu"
     repeat_region   23473..23761
                     /rpt_family="Alu"
     repeat_region   23744..23767
                     /note="(A)24"
     repeat_region   complement(24639..24694)
                     /rpt_type=tandem
                     /rpt_unit=A
     repeat_region   complement(25349..25713)
                     /rpt_family="MER42"
     repeat_region   complement(25727..26471)
                     /rpt_family="Alu"
     repeat_region   27191..27477
                     /rpt_family="Alu"
     repeat_region   27774..28057
                     /rpt_family="Alu"
     repeat_region   28040..28066
                     /note="(A)27"
     repeat_region   complement(28769..28838)
                     /rpt_type=tandem
                     /rpt_family="MER20"
     misc_feature    28987..29214
                     /note="GRAIL 2 excellent exon, frame 0"
     repeat_region   28712..28930
                     /rpt_family="Alu"
     misc_feature    join(29485..29559,29598..29779)
                     /note="98% identity dbSTS:G14522 (SHGC-11312)"
     misc_feature    29495..29976
                     /standard_name="AI027942"
                     /note="100% identity EST ov84a10.x1"
     misc_feature    complement(30401..30536)
                     /note="GRAIL 2 excellent exon, frame 2"
     repeat_region   complement(30682..30733)
                     /rpt_family="MIR"
     misc_feature    complement(31573..31724)
                     /note="GRAIL 2 excellent exon, frame 1"
     misc_feature    complement(32159..32232)
                     /note="GRAIL 2 excellent exon, frame 2"
     repeat_region   32388..32488
                     /rpt_family="MLT1"
     repeat_region   32617..32908
                     /rpt_family="MLT1"
     repeat_region   32977..33088
                     /rpt_family="Alu"
     repeat_region   complement(33670..33785)
                     /rpt_family="MLT1"
     misc_feature    complement(34021..34144)
                     /note="GRAIL 2 excellent exon, frame 2"
     misc_feature    complement(35238..35331)
                     /note="GRAIL 2 excellent exon, frame 2"
     repeat_region   36392..36663
                     /rpt_family="Alu"
     misc_feature    complement(36901..37222)
                     /rpt_family="Alu"
                     /note="100% identity EST ou55c09.x1"
```

```
     STS             /db_xref="dbEST:AI025011"
                     36901..37164
                     /standard_name="STSG-9983"
                     /db_xref="dbSTS:G26554"
     misc_feature    complement(37595..37654)
                     /note="GRAIL 2 excellent exon, frame 0"
     misc_feature    complement(join(38063..38218,38462..38578,38741..38995,
                     39071..39205,39532..39630,39935..40048,40300..40410,
                     40503..40661,41868..41972,42103..42225,42492..42569,
                     44379..44507))
                     /standard_name="histidyl-tRNA synthetase"
                     /note="78+-100% protein identity GenPept:U18937"
     misc_feature    complement(38069..38215)
                     /note="GRAIL 2 excellent exon, frame 0"

Note: remainder of annotations omitted.

...

         Query Match     1.2%;  Score 35;  DB 31;  Length 74371;
         Best Local Similarity 14.2%;  Pred. No. 3.7e-04;
         Matches   45;  Conservative  139;  Mismatches  131;  Indels   3;  Gaps   3;

Db    15940 RRAAAWTKSMRRMMWWWTWWWWRRAAAAWTTTWARKRAAAWYCAYAYWTTTTTTTTT 15899
               : |:|: |  |: :: ::||||  || :|||  || :|||| |||||
Qy     1008 gatattggttagagcaatcttcaagagaaaccgtaaatcaacctcaaaaacgtttgagc 1067

Db    15900 TTTTTTKWGASRSRGWSTCKCKCTSTSRCSCMSRSKSKRGWGYRSWKKYRCAMWMTCKSSK 15959
               : |:: | :|  ||||  | |:::: ::: ::::  :::::
Qy     1068 ctcccaacagtcccccttcccctccccatattacggtcaagcatgtgttcccaatcgtc- 1126

Db    15960 CWCWSYRMRMRKCYSCSYCYCSSGKKIWCRCSMYWYICYSYKYIYWSMSYCICTSWGWRWM 16019
             |:|:|:|:| || |:|| :|:|  |  |||||:|| :::|||||:|| || :|::||||||
Qy     1127 ctctgcgcatgctgttcccacgtgtatttctccgcgcgcagttctgttggag-gagga 1185

Db    16020 WSKGRSWMYASRSGCSCSCSMCMMCRCSCMSNKMWMWTTTTTKTRTWTTTWKRKAGASA 16079
             :::| |  : ::| | : ::|  | ||| ::::: :::::::|:|:|: ::::: | |:
Qy     1186 aggcgggacgttggcaactagcgctggctggagagttatatggccatccgaatcagtt-gtgg 1244

Db    16080 SRGKSKCRCMSYGKKSTCKMKMTCYCYMYSWSYKMKYSRYCCGCCYSSSYCKSCCWMMS 16139
             | | :  :| ||: : :||: :| :|: |:: : ::: :|||| | :::::| ::|
Qy     1245 cagtcaaatgctttgtctttttggccgctcatgtgagtatatcattctgtgaaactatct 1304

Db    16140 WRMKKGSWKKRMWWRCRW 16157
             |:|::| ::: :|||||
Qy     1305 aaatcaatatatgaatgtat 1322

RESULT 15

LOCUS       I28278       215 bp    DNA             PAT       30-OCT-1996
DEFINITION  Sequence 5 from patent US 5569830.
ACCESSION   I28278
NID         g1819054
VERSION     I28278.1  GI:1819054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 215)
  AUTHORS   Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
  TITLE     Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
  JOURNAL   Patent: US 5569830-A 5 29-OCT-1996;
FEATURES             Location/Qualifiers
     source          1..215
                     /organism="unknown"
BASE COUNT       15 a      8 c     25 g     26 t     141 others
ORIGIN

Db       10 SVVSRIASCNDKAKKDGNTTSSWTDCCNRIWGVCDTDTYRVNNDSGHNKYSSANYNYG 69

         Query Match     1.1%;  Score 30;  DB 25;  Length 215;
         Best Local Similarity 17.0%;  Pred. No. 1.48e-01;
         Matches   31;  Conservative  70;  Mismatches  78;  Indels   3;  Gaps   3;
```

```
Qy    843  catgataggtttaataataaatatataaaatcttatattttagtaagtctagtcaaactta  902

Db     70  GNNVGAAKT-HYYTHTNVSGADSKTVTDSYNASGTSSSNGGTDGNRSGGADSYGSSKTPAMT  128
           : :| ::  |:|:::||:::|:    | ||:|::  |:::||| :::| ::||:  :||
Qy    903  agagctttgatgtcgcacatagttgtttaaacaaggtgttcttcc-atgttcgtt-ct  960

Db    129  SRNRTGKTANNAVDSRNMGDASVGSDKNTKHAKNSADGKVGSKNNGDRNNRYGTGTKSN  188
           : :||:     |||: : ::||::  |::  :| ||:
Qy    961  aatatgtggatagtattccgattcattcgccagaggtgtggcctgtggatatttggttag  1020

Db    189  vs 190
           ::
Qy   1021  ag 1022
```